

A Study on the Short-term Load Forecasting using Support Vector Machine

趙南熏<sup>†</sup> · 宋敬彬<sup>\*</sup> · 盧永秀<sup>\*</sup> · 姜大昇<sup>\*</sup>

(Nam-Hoon Jo · Kyung-Bin Song · Youngsu Roh · Daeseung Kang)

**Abstract** - Support Vector Machine(SVM), of which the foundations have been developed by Vapnik (1995), is gaining popularity thanks to many attractive features and promising empirical performance. In this paper, we propose a new short-term load forecasting technique based on SVM. We discuss the input vector selection of SVM for load forecasting and analyze the prediction performance for various SVM parameters such as kernel function, cost coefficient C, and  $\epsilon$  (the width of  $\epsilon$ -tube). The computer simulation shows that the prediction performance of the proposed method is superior to that of the conventional neural networks.

**Key Words** : Load Forecasting, Support Vector Machine, Nonlinear Regression, Kernel Function

1. 서 론

전력 수요 예측은 전기공학 분야에서 매우 활발히 진행되고 있는 연구 분야이다. 특히, 전력시장이 경쟁체제로 바뀌어 가면서 단기 전력 수요 예측은 그 중요성이 급속히 증가되고 있다. 그러나, 전력 수요 예측은 매우 어려운 것으로 알려져 있으며, 이를 위한 많은 연구가 전 세계적으로 수행되었다. 이러한 연구는 크게 두 가지, 시계열 모델(time series model)과 인과 모델(causal model)로 분류될 수 있는데, 전자에서는 전력 수요가 이전 수요의 함수로 주어지는 것이며, 후자에서는 전력 수요가 날씨나 사회적 이벤트 등과 같은 외부 요소의 함수로 주어진다. 시계열 모델과 관련된 연구결과에는 동적 선형 모델 [1], 문턱 자기회귀 모델(threshold autoregressive model) [2], Kalman filter를 이용한 예측기법 [3], [4] 등이 있다. 인과모델 관련 연구는 ARMAX 모델 [5], 비매개변수화 회귀(nonparametric regression) [6], Curve-fitting 절차 [7] 등이 있다. 최근에는 인공지능기법을 적용한 전력 수요 예측기법 연구가 매우 활발히 진행되고 있다. 전문가 시스템을 이용한 기법 [8], 퍼지 추론(Fuzzy inference)기법 [9], 퍼지-뉴로(Fuzzy -neural)기법 [10], 퍼지선형회귀법 [11] 등이 연구되어 왔다. 이러한 인공지능기법 중 가장 많은 관심을 받아온 기법은 1990년대 초부터 연구되어 온 신경회로망(Neural Networks; NN)이다 [12], [13], [14], [15]. 그러나, 신경회로망에는 몇 가지 단점이 지적되고 있다. 첫째로 학습에 필요한 시간이 오래 소요

되고, 둘째로 Overfitting시 예측성능이 저하되며, 셋째로 최적의 설계변수를 결정할 때 국소적 최소점(local minimum)으로 수렴하는 것 등이 취약한 점으로 알려져 있다.

본 논문에서는 최근에 인공지능의 새로운 분야로 각광받고 있는 지원벡터머신(Support Vector Machine; SVM)을 사용하여 기존의 예측기법보다 우수한 성능을 갖는 예측기법을 제시하고자 한다. SVM은 기존의 통계적 학습 방법에서 이용되는 '경험적 위험 최소화' (Empirical Risk Minimization; ERM)와는 다른 '구조적 위험 최소화' (Structural Risk Minimization; SRM)을 이용하여 일반화 오차를 감소시키는 기법이다. ERM이 학습데이터의 오차를 최소화하는 것과는 대조적으로, SRM은 기대위험(expected risk)의 상한을 최소화하는 것이다. SVM이 기존의 신경회로망보다 우수한 일반화 성능을 갖는 것은 바로 이러한 때문이다. SVM은 초기에 패턴인식과 같은 분류(classification) 문제에 적용되어 왔으나, 최근에 비선형 회귀(nonlinear regression) 문제에도 적용되기 시작하였다 [16], [17], [18], [19], [20], [21]. 본 논문에서는 SVM의 기본 이론을 분석한 후, SVM을 이용한 전력수요예측기법이 기존 기법보다 우수한 예측성능을 보이는 것을 모의실험을 통하여 보이고자 한다.

본 논문의 구성은 다음과 같다. 2장에서는 선형 지원벡터머신(SVM), 쌍대문제(Dual problem)등을 우선 소개하고, 이를 비선형 SVM으로 확장하기 위한 커널기교(kernel trick) 기법을 알아본다. 3장에서는 SVM을 이용한 전력수요예측기법을 제시한다. 즉, SVM의 입력벡터에 사용되는 특징선택(feature selection), 자료분할, 입력벡터 크기조정(scaling) 등을 소개한다. 4장에서는 제시된 기법의 우수성을 검증하기 위한 모의실험 결과를 소개한다. 모의실험 결과, 제시된 예측기법은 기존의 신경회로망 예측기법보다 우수한 예측성능을 보임을 알 수 있었다. 마지막으로 5장에서는 본 논문의 결론을 제시한다.

<sup>†</sup> 교신저자, 正會員 : 崇實大 工大 電氣工學部 助教授 · 工博  
E-mail : nhjo@ssu.ac.kr

<sup>\*</sup> 正會員 : 崇實大 工大 電氣工學部 助教授 · 工博  
接受日字 : 2006年 3月 28日  
最終完了 : 2006年 5月 23日

## 2. 지원벡터머신(Support Vector Machine; SVM)

### 2.1 선형 지원벡터회귀(Support Vector Regression)

#### 2.1.1 지원벡터회귀 기본 개념

선형함수

$$f(x) = \langle w, x \rangle + b \quad (1)$$

를 사용하여 데이터 집합

$$D = \{(x_1, y_1), \dots, (x_l, y_l)\}, \quad x \in R^n, \quad y \in R$$

을 근사하는 문제를 고려하자.  $\epsilon$ -SV regression의 목적은 모든 학습 데이터(training data)의 최대 편차가  $\epsilon$  이내인 함수  $f(x)$  중에서, 최대로 편평(flat)한 함수  $f(x)$ 를 구하는 것이다 (여기서 편차란 학습 데이터의 목표치(target)  $y_i$ 와  $f(x)$ 를 통한 학습 데이터의 사상  $f(x_i)$ 의 차이를 의미한다). (1)에서의 편평함(flatness)은  $w$ 의 노름(norm)을 최소화하는 것을 의미하므로,  $\epsilon$ -SV regression는 볼록 최적화(convex optimization) 문제

$$\min_{w, b} \frac{1}{2} \|w\|^2 \quad (2)$$

$$\text{subject to } \begin{cases} y_i - \langle w, x_i \rangle - b \leq \epsilon \\ \langle w, x_i \rangle + b - y_i \leq \epsilon \end{cases}$$

로 표현할 수 있다. 주어진 데이터 집합  $D$ 에 의하여, 볼록 최적화문제 (2)를 만족하는 함수  $f(x)$ 는 존재할 수도 있고 존재하지 않을 수도 있다. 함수  $f(x)$ 가 존재하지 않는 경우에 볼록 최적화 문제 (2)를 풀기 위해서는 약간의 오차를 인정해야 한다. 이를 위해서 슬랙(slack) 변수  $\xi_i, \xi_i^*$ 를 도입하여 Vapnik이 제시한 다음과 같은 문제를 고려할 수 있다:

$$\min_{w, b, \xi, \xi^*} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l (\xi_i + \xi_i^*) \quad (3)$$

$$\text{subject to } \begin{cases} y_i - \langle w, x_i \rangle - b \leq \epsilon + \xi_i \\ \langle w, x_i \rangle + b - y_i \leq \epsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0, i = 1, \dots, l \end{cases}$$

여기서, 상수  $C$ 는  $\epsilon$ 보다 큰 오차를 얼마만큼 허용할 것인가와 함수  $f(x)$ 의 편평함(flatness) 사이의 타협점을 결정하는 것으로, 비용계수(Cost Coefficient)라고 부른다. 비용계수  $C$ 가 상대적으로 큰 경우에는 학습데이터의 오차를 가능한 작게 만드는 것이고, 반대로  $C$ 가 상대적으로 작은 경우에는  $w$ 의 노름(norm)을 작게 만들기 때문에 함수  $f(x)$ 를 가능한 편평하게 만드는 것이다. 최적화문제 (3)은 쌍대(dual) 문제로 바꾸어 고려할 경우, 훨씬 간단히 문제가 풀린다는 것이 밝혀져 있다. 또한, 쌍대문제의 경우, SVM을 비선형 함수로 확장하는데 매우 용이하다는 것도 알려져 있다.

#### 2.1.2 쌍대(Dual) 문제 설정

쌍대문제를 만들기 위하여 우선 목적함수와 제약조건으로부터 라그랑주(Lagrange) 함수를 아래와 같이 정의한다:

$$L = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l (\xi_i + \xi_i^*) - \sum_{i=1}^l \alpha_i (\epsilon + \xi_i - y_i + \langle w, x_i \rangle + b) - \sum_{i=1}^l \alpha_i^* (\epsilon + \xi_i^* + y_i - \langle w, x_i \rangle - b) - \sum_{i=1}^l (\eta_i \xi_i + \eta_i^* \xi_i^*) \quad (4)$$

(4)식에서 쌍대변수(dual variable)는 조건

$$\alpha_i, \alpha_i^*, \eta_i, \eta_i^* \geq 0$$

을 만족하여야 하며,  $(w, b, \xi_i, \xi_i^*)$ 에 대한  $L$ 의 편미분은 0이어야 한다:

$$\begin{aligned} \partial_b L &= \sum_{i=1}^l (\alpha_i^* - \alpha_i) = 0 \\ \partial_w L &= w - \sum_{i=1}^l (\alpha_i - \alpha_i^*) x_i = 0 \\ \partial_{\xi_i} L &= C - \alpha_i^* - \eta_i = 0 \end{aligned} \quad (5)$$

(5)을 (4)에 대입하면 다음과 같은 (3)의 쌍대 최적화(Dual optimization) 문제를 구할 수 있다:

$$\max \begin{cases} -\frac{1}{2} \sum_{i,j=1}^l (\alpha_i - \alpha_j)(\alpha_i - \alpha_j) \langle x_i, x_j \rangle \\ -\epsilon \sum_{i=1}^l (\alpha_i + \alpha_i^*) + \sum_{i=1}^l y_i (\alpha_i - \alpha_i^*) \end{cases}$$

$$\text{subject to } \begin{cases} \sum_{i=1}^l (\alpha_i - \alpha_i^*) = 0 \\ \alpha_i, \alpha_i^* \in [0, C] \end{cases} \quad (6)$$

(6)을 유도하면서 쌍대변수  $\eta_i, \eta_i^*$ 는 쌍대 목적함수에는 더 이상 나타나지 않고 단지 쌍대 실행조건(feasibility condition)에만 나타나기 때문에, (5)의 세 번째 조건을 사용하여 소거하였다. 조건 (5)의 두 번째 조건을 다시 쓰면

$$w = \sum_{i=1}^l (\alpha_i - \alpha_i^*) x_i$$

이 되고, 이로 부터

$$f(x) = \sum_{i=1}^l (\alpha_i - \alpha_i^*) \langle x_i, x \rangle + b \quad (7)$$

을 얻을 수 있다.

#### 2.1.3 상수항 b의 계산

함수  $f(x)$ 의  $b$ 를 계산하기 위해서는 Karush-Kuhn-Tucker (KKT) 조건을 이용한다. 이는 쌍대변수와 제약조건이 최적해에서는 항상 0이 된다는 것이다. 따라서

$$\begin{aligned} \alpha_i (\epsilon + \xi_i - y_i + \langle w, x_i \rangle + b) &= 0 \\ \alpha_i^* (\epsilon + \xi_i^* + y_i - \langle w, x_i \rangle - b) &= 0 \end{aligned} \quad (8)$$

와

$$\begin{aligned} (C - \alpha_i) \xi_i &= 0 \\ (C - \alpha_i^*) \xi_i^* &= 0 \end{aligned} \quad (9)$$

을 얻을 수 있는데, 이로부터 몇 가지 중요한 점을 알 수 있다. 첫 번째로,  $\alpha_i^* = C$ 에 해당하는 샘플  $(x_i, y_i)$  들만  $f(x)$  주위의  $\epsilon$ -비민감 튜브( $\epsilon$ -insensitive tube)의 바깥쪽

에 존재할 수 있다는 것을 (9)로부터 알 수 있다. 두 번째로,  $a_i a_i^* = 0$ 을 얻을 수 있다. 즉, 쌍대변수  $a_i, a_i^*$  중 최소한 한개는 0이라는 것이다. 왜냐하면, 그렇지 않은 경우 0이 아닌 슬랙 변수가 양방향으로 존재한다는 것인데, 이는 해당 데이터가  $\epsilon$ -튜브의 위와 아래에 동시에 존재함을 의미하여 불가능하기 때문이다. 세 번째로는,  $a_i^{(*)} \in (0, C)$ 인 경우에는, (9)으로부터  $\xi_i^{(*)} = 0$ 을 얻을 수 있으며, 또한, (8)식의 두 번째 인자가 0이 되어야만 함을 알 수 있다. 따라서,  $b$ 는 아래와 같이 계산할 수 있다 :

$$b = \begin{cases} y_i - \langle w, x_i \rangle - \epsilon & \text{for } a_i \in (0, C) \\ y_i - \langle w, x_i \rangle + \epsilon & \text{for } a_i^* \in (0, C) \end{cases} \quad (10)$$

### 2.2 비선형 지원벡터회귀

앞서 살펴본 선형 SVM을 비선형으로 확장하기 위하여 커널함수(Kernel function)를 이용하는데, 이를 커널 기교(Kernel trick)이라고 한다. 커널 기교란 어떤 사상(mapping)  $\Phi(x)$ 를 통하여 원래의 데이터 공간을 고차원의 새로운 공간<sup>1)</sup>으로 사상시킨 후, 이 새로운 공간에서 앞서 살펴본 선형 SVM regression 이론을 적용하는 것이다. 앞서의 선형 SVM의 중요한 식인 (6)을 다시 살펴보면, 데이터들은 항상 내적(dot product)의 형태로만 SV 알고리즘에 나타나는 것을 알 수 있다. 따라서,  $x$ 를  $\Phi(x)$ 로 교체하는 커널기교를 사용할 경우, 수정되는 부분은  $\langle x_i, x_j \rangle$  뿐이며, 이는  $\langle \Phi(x_i), \Phi(x_j) \rangle$ 로 바뀌게 된다는 것을 알 수 있다. 따라서, 커널 함수를

$$k(x_1, x_2) = \langle \Phi(x_1), \Phi(x_2) \rangle$$

로 정의하면 (6)로부터 비선형 SVM 최적화문제가 다음과 같이 표현됨을 알 수 있다 :

$$\begin{aligned} \max \quad & \left\{ \begin{aligned} & -\frac{1}{2} \sum_{i,j=1}^n (a_i - a_i^*)(a_j - a_j^*) k(x_i, x_j) \\ & -\epsilon \sum_{i=1}^n (a_i + a_i^*) + \sum_{i=1}^n y_i (a_i - a_i^*) \end{aligned} \right. \\ \text{subject to} \quad & \begin{cases} \sum_{i=1}^n (a_i - a_i^*) = 0 \\ a_i, a_i^* \in [0, C] \end{cases} \end{aligned} \quad (11)$$

(11)식으로부터 커널 함수를 사용하는 경우, 함수  $\Phi(x)$ 를 명시적으로 구할 필요가 없음을 알 수 있다. 또한,

$$w = \sum_{i=1}^n (a_i - a_i^*) \Phi(x_i)$$

을 얻을 수 있으며, (7)로부터 지원벡터전개는

$$f(x) = \sum_{i=1}^n (a_i - a_i^*) k(x_i, x) + b \quad (12)$$

로 바뀌는 것을 알 수 있다.

1) 이를 특징공간(Feature space)이라고 한다.

### 3. SVM을 이용한 수요예측

#### 3.1 특징 선택

학습 데이터의 각각의 원소들을 특징(feature)이라고 부르는데, 이러한 특징의 선택은 예측성능에 많은 영향을 미치게 된다. 이러한 특징으로 선택할 수 있는 정보 몇 가지를 소개하기로 한다.

##### 3.1.1 요일

전력 수요는 주별로 주기적인 성질이 있다. 또한, 휴일의 전력 수요는 평일의 전력 수요보다 일반적으로 낮다. 따라서, 요일 정보와 휴일정보를 학습데이터에 포함시키는 것은 당연하다. 실제로 대부분의 많은 연구에서도 요일 정보와 휴일정보를 학습데이터의 특징을 선택하여 왔다.

##### 3.1.2 온도

여름의 전력 수요를 예측하는데 중요한 정보중의 하나가 예측일 또는 그 이전일의 온도이다. 온도가 높을수록 냉방기의 사용이 증가하여 전력을 많이 사용하는 요인이 되는 것이다. 따라서, 온도 정보를 이용하면 예측성능을 향상시키는데 많은 도움을 줄 것으로 예상할 수 있다. 하지만, 온도를 학습데이터에 포함시키는 것은 현실적인 어려움이 있다. 예측일의 온도의 정확한 값은 알 수가 없으며, 설정 온도의 예측치를 사용한다고 하더라도 예측온도 자체가 오차를 필연적으로 포함하고 있기 때문에 신뢰할 수 없다는 것이다. 하지만, 온도를 예측하는 것 자체가 매우 어려운 문제이기 때문에 온도를 학습데이터로 사용하는 것은 어려운 문제이다.

##### 3.1.3 이전 전력수요

학습데이터에 포함되어야 하는 또 하나의 중요한 것은 이전 전력수요로서, 이는 시계열(time-series) 개념을 SVM에 포함시키는 것이다. 다시 말해서,  $y_i$ 가 예측을 위한 목표치라면, 입력벡터  $x_i$ 가 몇 개의 이전 전력수요들  $y_{i-1}, y_{i-2}, \dots, y_{i-\Delta}$ 을 포함하도록 하는 것이다. 예를 들어  $\Delta = 7$ 이라면, 2002년 7월 1일의 전력수요를 예측하기 위해서는 2002년 6월 24일에서 30일까지의 전력수요 자체를 입력벡터  $x_i$ 에 포함시켜야 한다는 것이다. 일반적으로  $\Delta = 1$  또는  $\Delta = 7$  등이 많이 사용된다.

#### 3.2 자료 분할

전력 수요 패턴은 계절별로 비슷한 형태를 보이기 때문에, 주어진 데이터 중 일부분만을 사용할 경우 더 좋은 예측성능을 보일 가능성이 있다. 본 연구에서는 7월과 8월의 전력수요를 예측하고자 하는 것이므로, 학습데이터도 여름 데이터를 사용하는 것이 우수한 예측성능을 보일 것으로 판단된다. 따라서, 본 모의실험에서는 주어진 학습데이터 중 여름과 비슷한 패턴을 보일 것으로 예상되는 6월, 7월, 8월, 9월의 데이터만을 사용하여 SVM을 학습하였다.

### 3.3 입력벡터 크기조정

입력벡터 크기조정은 매우 중요한 과정으로서 크게 두 가지 장점이 있다. 첫 번째로는 계산과정 중 일어나는 수치적인 문제를 회피하는 것이 가능하다. 예를 들어, 다항식 커널의 경우, 수치적으로 매우 큰 값을 갖는 속성이 수치적인 문제를 야기할 수 있는데 이를 방지할 수 있다. 두 번째로, 수치적으로 매우 큰 값을 갖는 속성(attribute)에 의해서 작은 값을 갖는 속성이 실질적으로 무시되는 현상을 방지할 수 있다. 실제로, 전력수요의 경우 양의 값을 가지며 그 크기도 천만 kW 단위이지만, 온도의 범위는 -10°C에서 40°C 정도의 값을 갖는다. 이와 같은 이유로, 입력벡터 크기조정 과정을 수행하지 않으면 알고리즘이 수렴하지 않고 발산하거나, 수렴한다고 하더라도 성능이 별로 좋지 않은 경우가 대부분이다.

그러므로, 본 연구에서는 입력벡터의 모든 성분 변수가 0에서 1사이의 값이 되도록 크기조정하여 SVM을 학습시킨다. 여기서 주의할 점은, 학습된 SVM을 이용하여 전력수요를 예측할 때 사용하는 입력벡터도 반드시 위와 같은 크기조정을 거친 후 사용하여야 한다는 것이다.

## 4. 모의실험

### 4.1 예측기간 및 MAPE

본 연구에서 제안된 수요예측 알고리즘의 성능을 분석하기 위하여 다음 기간의 하루 최대 전력수요량을 예측하여 보았다.

- ① 2002년 여름 (7월, 8월)
- ② 2003년 여름 (7월, 8월)

이 기간들은 제한절과 광복절과 같은 특수일을 포함하고 있는데, 이러한 특수일에는 예측성능이 일반적으로 저하되는 특징이 있으므로, 예측알고리즘의 성능 분석에 효과적일 것으로 예상된다. 또한, 7월 마지막 주와 8월 첫째 주 기간은 일반적으로 휴가기간이다. 이 기간 동안은 예측오차가 상대적으로 큰 것을 볼 수 있는데, 이는 산업체의 생산 활동 중단으로 인해 전력수요가 감소하기 때문이다.

예측오차를 계산하는 데에는 퍼센트 절대값 평균오차 (Mean Absolute Percentage Error; MAPE)를 사용하였다.

$$MAPE = 100 \times \frac{1}{n} \times \sum_{i=1}^n \left| \frac{L_i - \hat{L}_i}{L_i} \right|,$$

여기서,  $L_i$ 와  $\hat{L}_i$ 는  $i$  번째 날의 실제 최대 전력수요와 예측된 최대 전력수요이며,  $n$ 은 예측기간내의 일수이다.

### 4.2 기존의 신경회로망(NN)과의 예측성능 비교

본 연구에서 제안하는 SVM을 이용한 전력 수요 예측 알고리즘의 우수성을 입증하기 위해서는 우선 기존의 예측 알고리즘과의 비교가 필요하다. 최근에 신경회로망을 이용한 국내의 전력 수요 예측 연구결과가 [22]에서 발표되었다. [22]에서는 입력값 선정과 학습 케이스 선정 등과 같이

예측정확성에 영향을 주는 요인에 대해 과거 논문을 조사하고 평가하여 최소구조 신경회로망을 제안하였는데, 기존의 연구결과 [23]에 비해서 정확도가 상당히 향상되었다. 따라서, 본 논문에서 제안된 기법의 우수성을 입증하는데 [22]의 결과와만 비교해도 충분할 것으로 판단된다. [22]에서는 신경회로망의 입력벡터의 형태로 아래와 같은 형태를 사용하였다.

$$[L(t-1), T_{\max}(t), T_{\min}(t), T_{\max}(t-1), T_{\min}(t-1)]$$

여기서,  $L(t-1),$

$T_{\max}(t), T_{\min}(t), T_{\max}(t-1), T_{\min}(t-1)$ 은 각각 예측일 하루 전일의 최대 수요, 예측일의 최고온도, 예측일의 최저온도, 예측일 하루 전일의 최고온도, 예측일 하루 전일의 최저온도를 나타낸다. 또한, [22]에서는 요일정보를 고려하기 위하여 다음과 같은 네 가지 경우별로 신경회로망 학습을 별도로 수행하였다.

- ① 평일(화~금)
- ② 토요일
- ③ 일요일
- ④ 월요일

[22]와 공평하게 비교하기 위하여 SVM의 입력벡터에 사용하는 정보는 예측일 하루 전일의 최대 수요, 예측일의 특수일 여부, 예측일의 최고온도, 예측일의 최저온도, 예측일 하루 전일의 최고온도, 예측일 하루 전일의 최저온도로 제한하였다. 이를 이용한 SVM의 입력벡터는 다음과 같은 형태를 사용하였다.

[이전수요, 요일, 온도]

여기서 이전수요는 예측일 하루 전일의 최대 수요이며, 온도는 네 가지 온도, 즉 예측일의 최고온도, 예측일의 최저온도, 예측일 하루 전일의 최고온도, 예측일 하루 전일의 최저온도를 나타낸다. 요일은 요일정보와 특수일 여부를 나타내기 위한 7자리 이진 코드인데, 앞의 6자리는 요일을 나타내며 마지막 7번째 이진 코드는 특수일 여부를 나타낸다. 예를 들어 특수일이 아닌 월요일, 화요일, 일요일등은 각각 [1 0 0 0 0 0 0], [0 1 0 0 0 0 0], [0 0 0 0 0 0 0]으로 나타내고, 화요일인데 특수일인 경우에는 [0 1 0 0 0 0 1]으로 나타낸다.

SVM을 학습시키는데 필요한 커널함수는 다음과 같은 가우시안 방사기저함수(RBF) 커널을 선택하였다:

$$k(x_1, x_2) = \exp\left(-\frac{\|x_1 - x_2\|^2}{2\sigma^2}\right)$$

또한,  $\epsilon$ 은 일반적으로 적을수록 우수한 일반화성능을 나타내기 때문에  $\epsilon = 0.5$ 를 선택하였다. 또한, (3)으로부터 비용계수  $C$ 는 충분히 큰 값을 사용하여야 슬랙변수에 의한 오차를 줄일 수 있음을 알 수 있으며, 본 논문에서는  $C = 10^7$ 을 사용하였다. 마지막으로 커널함수의 대역폭 (bandwidth)는  $\sigma = 5, 10, 20, 30$ 인 경우에 대해서 모의실험을 수행하였다. 한편, SVM을 학습하는데 주어진 학습 데이터는 다음과 같다:

- ① 2000년 1월 1일 - 2002년 6월 31일 : 2002년 여름 (7월, 8월) 예측용
- ② 2000년 1월 1일 - 2003년 6월 31일 : 2003년 여름 (7월, 8월) 예측용

앞서 기술한 바와 같이 SVM 학습시 자료 분할을 수행하며, 여기서는 실제 학습데이터로 다음을 사용하였다.

- ① 2002년 여름 (7월, 8월) 예측시 사용한 학습데이터
  - 2000년 6월, 7월, 8월, 9월
  - 2001년 6월, 7월, 8월, 9월
  - 2002년 6월
- ② 2003년 여름 (7월, 8월) 예측시 사용한 학습데이터
  - 2000년 6월, 7월, 8월, 9월
  - 2001년 6월, 7월, 8월, 9월
  - 2002년 6월, 7월, 8월, 9월
  - 2003년 6월

표 1. SVM기법과 신경회로망기법의 예측성능 비교  
Table 1. Comparison of prediction performance between SVM and Neural Network schemes

	2002년 (MAPE)	2003년 (MAPE)	두해 평균 (MAPE)
신경회로망[22]	3.20	2.71	2.96
SVM, $\sigma = 5$	3.15	2.32	2.74
SVM, $\sigma = 10$	2.97	2.24	2.61
SVM, $\sigma = 20$	2.93	2.25	2.59
SVM, $\sigma = 30$	2.95	2.25	2.60

MATLAB을 이용하여 위에서 언급한 방식대로 SVM을 학습시킨 후, SVM의 예측성능을 계산해 보았다. 표 1은 신경회로망 기법과 제안된 SVM기법 ( $\sigma = 5, 10, 20, 30$ )의 예측오차를 정리한 것이다. 표 1로부터,  $\sigma = 5, 10, 20, 30$ 에 대해서, SVM 기법이 기존의 신경회로망 기법보다 우

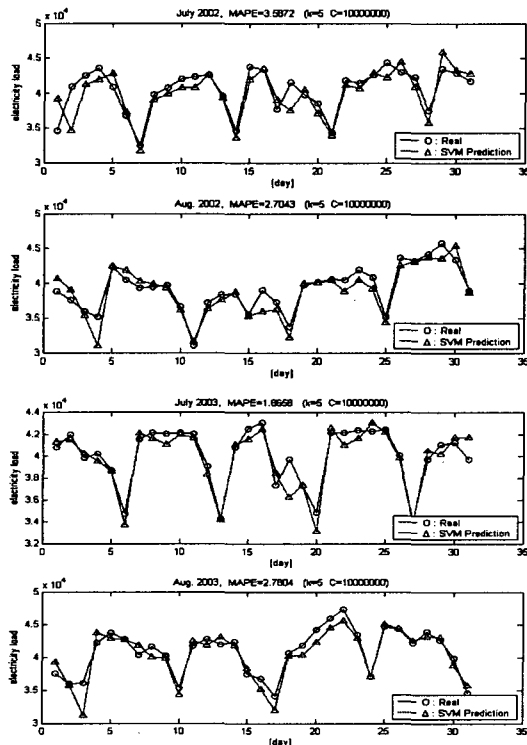


그림 1.  $\sigma = 5, C = 10^7$ 일 때의 SVM 예측결과  
Fig. 1 SVM prediction result when  $\sigma = 5, C = 10^7$

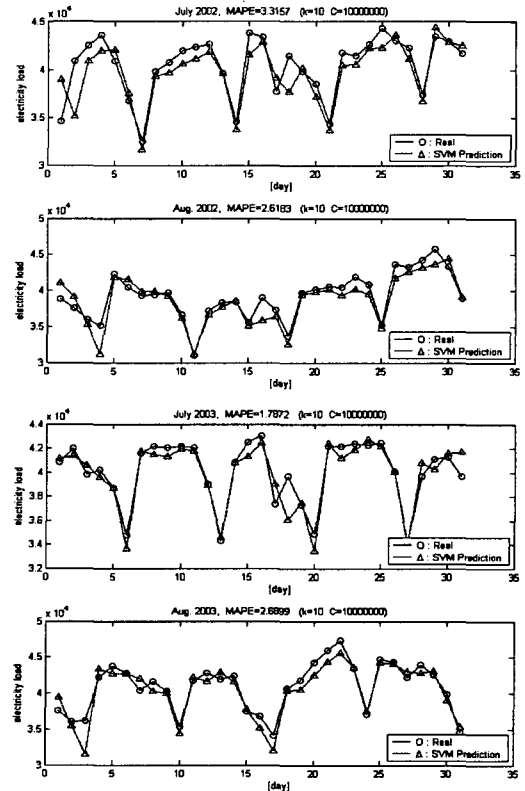


그림 2.  $\sigma = 10, C = 10^7$ 일 때의 SVM 예측결과  
Fig. 2 SVM prediction result when  $\sigma = 10, C = 10^7$

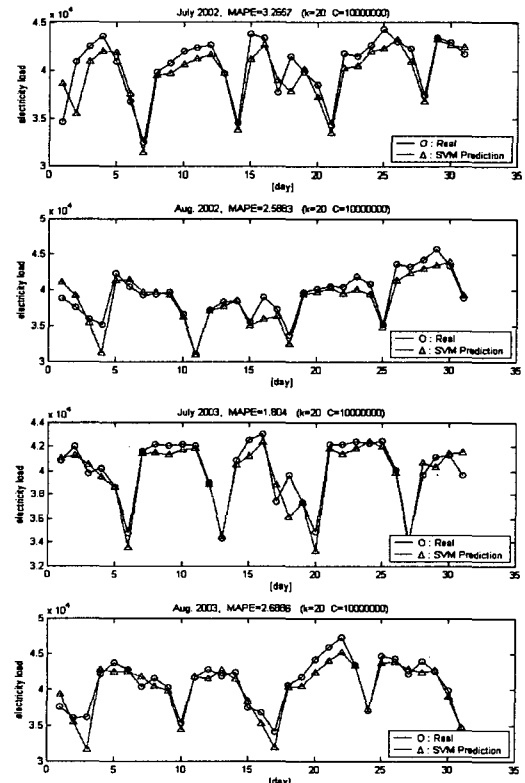


그림 3.  $\sigma = 20, C = 10^7$ 일 때의 SVM 예측결과  
Fig. 3 SVM prediction result when  $\sigma = 20, C = 10^7$

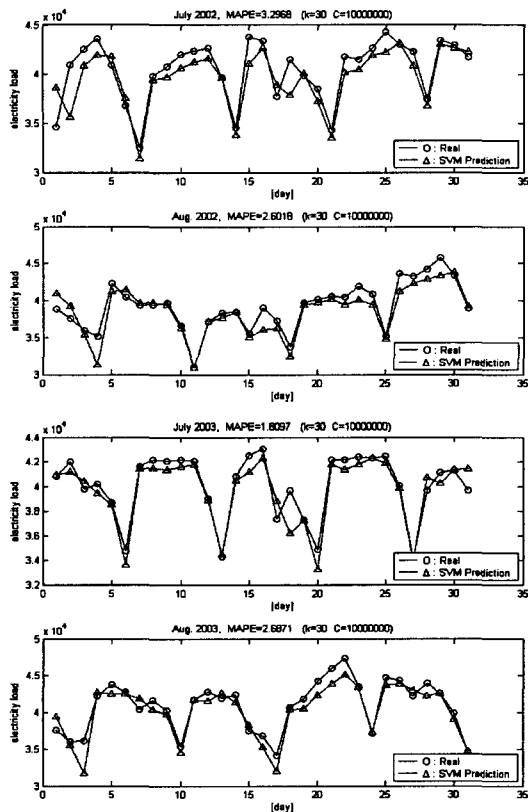


그림 4.  $\sigma = 30, C = 10^7$  일 때의 SVM 예측결과  
 Fig. 4 SVM prediction result when  $\sigma = 30, C = 10^7$

수한 예측성능을 갖고 있는 것을 확인할 수 있다. 특히,  $\sigma = 20$ 인 경우, SVM기법이 신경회로망 기법보다 예측성능이 12.8% 향상되었음을 알 수 있다. 그림 1, 2, 3, 4는 각각  $\sigma = 5, 10, 20, 30$  일 때의 예측결과를 보여준다.

### 5. 결 론

본 논문에서는 SVM을 이용한 단기전력수요예측기법을 제시하였다. 모의실험을 통하여 제안된 기법이 기존의 신경회로망 기법보다 우수한 예측성능을 보임을 입증하였다. SVM 기법은 여러 가지 응용분야에서 우수한 비선형회귀성능을 보이고 있기 때문에, 이는 이미 예측되었던 결과이다. 앞으로, 입력벡터의 변화 및 자료분할의 변화 등에 대해서도 예측성능을 계산하는 연구가 필요할 것으로 보이며, 특수일의 전력수요 예측에도 적용가능한 SVM 기법의 연구도 가능할 것으로 기대된다. 또한, 4장의 예측결과 그림에서 7월1일과 7월2일의 예측성능이 좋지 않은 것을 볼 수 있는데, 이는 본 논문에서 제시된 SVM기법만의 문제는 아닌 것으로 판단된다. 왜냐하면, 7월1일이 휴일이 아닌 월요일임에도 불구하고 전력수요가 다른 주의 월요일보다 상당히 작은 것을 확인할 수 있는데, 이러한 점이 예측성능을 악화시킨 것으로 판단된다. 이러한 날의 예측성능을 향상시키기 위해서는 특수일의 예측기법과 유사한 기법을 병행 사용해야 할 것으로 보이며, 이에 대한 연구가 현재 진행 중이다.

### 감사의 글

본 연구는 산업자원부의 지원에 의하여 기초전력연구원(R-2005-B-115)주관으로 수행된 과제임.

### 참 고 문 헌

- [1] A.P. Douglas, A.M. Breipohl, F.N. Lee, and R. Adapa, "The impact of temperature forecast uncertainty on bayesian load forecasting," IEEE Trans. Power Systems, vol. 13, no. 4, pp. 1507 - 1513, 1998.
- [2] S.R. Huang, "Short-term load forecasting using threshold autoregressive models," IEEE Proc.-Gener. Transm. Distrib., vol. 144, no. 5, pp. 477-481, 1997.
- [3] D.G. Infield and D.C. Hill, "Optimal smoothing for trend removal in short term electricity demand forecasting," IEEE Trans. Power Systems, vol. 13, no. 3, pp. 1115-1120, 1998.
- [4] J.H. Park, Y.M. Park, and K.Y. Lee, "Composite modeling for adaptive short-term load forecasting," IEEE Trans. Power Systems, vol. 6, no. 2, pp. 450 - 457, 1991.
- [5] H.T. Yang and C.M. Huang, "A new short-term load forecasting approach using self-organizing fuzzy ARMAX models," IEEE Trans. Power Systems, vol. 13, no. 1, pp. 217-225, 1998.
- [6] W. Charytoniuk, M.S. Chen, and P. Van Olinda, "Nonparametric regression based short-term load forecasting," IEEE Trans. Power Systems, vol. 13, no. 3, pp. 725-730, 1998.
- [7] J.W. Taylor and S. Majithia, "Using combined forecasts with changing weights for electricity demand profiling," J. Oper. Res. Soc., vol. 51, no. 1, pp. 72-82, 2000.
- [8] S. Rahman and O. Hazim, "A generalized knowledge-based short-term load-forecasting technique," IEEE Trans. Power Systems, vol. 8, no. 2, pp. 508-514, 1993.
- [9] H. Mori and H. Kobayashi, "Optimal fuzzy inference for short-term load forecasting," IEEE Trans. Power Systems, vol. 11, no. 1, pp. 390-396, 1996.
- [10] S.E. Papadakis, J.B. Theocharis, S.J. Kiartzis, and A.G. Bakirtzis, "A novel approach to short-term load forecasting using fuzzy neural networks," IEEE Trans. Power Systems, vol. 13, no. 2, pp. 480-492, 1998.
- [11] Kyung-Bin Song, Young-Sik Baek, Dug Hun Hong, and Jang, G., "Short-term load forecasting for the holidays using fuzzy linear regression method," IEEE Trans. Power Systems, vol. 20, no. 1, pp. 96-101, 2005.
- [12] I. Drezga and S. Rahman, "Input variable selection for ANN-based short-term load forecasting," IEEE

Trans. Power Systems, vol. 13, no. 4, pp. 1238-1244, 1998.

- [13] H.S. Hippert, C.E. Pedriera, and R.C. Souza, "Neural Networks for Short-term Load Forecasting: A Review and Evaluation," IEEE Trans. Power Systems, vol. 16, no. 1, pp. 44-55, 2001.
- [14] A. Khotanzad, R. Afkhani-Rohani, and D. Maratukulam, "ANNSTLF - Artificial neural network short-term load forecaster - Generation three," IEEE Trans. Power Systems, vol. 13, no. 4, pp. 1413-1422, 1998.
- [15] Reis, A.J.R. and da Silva, A.P.A., "Feature extraction via multiresolution analysis for short-term load forecasting," IEEE Trans. Power Systems, vol. 20, no. 1, pp. 189-198, 2005.
- [16] C.H.C. Burges, "A tutorial on Support Vector Machines for Pattern Recognition," Data Mining and Knowledge Discovery," pp. 121-167, 1998.
- [17] B.J. Chen, M.W. Chang, and C.J. Lin, "Load Forecasting using SVM : A study on EUNITE Competition 2001," IEEE Trans. Power Systems, 2004, pp. 1821 - 1830.
- [18] N. Cristianini and J. Shawe-Taylor, "An introduction to Support Vector Machines," Cambridge University press, 2003.
- [19] S.R. Gunn, "Support Vector Machines for Classification and Regression," Technical Report, 1998.
- [20] A.J. Smola, B. Scholkopf, "A tutorial on support vector regression," Neuro COLT2 Technical Report Series, 1998.
- [21] V. Vapnik, S.E. Golowich, A.Smola, "Support Vector Method for function approximation, regression estimation, and signal processing," Advances in neural information processing systems, pp. 281-287, 1997.
- [22] 하성관, 송경빈, 김홍래, "신경회로망과 하절기 온도 민감도를 이용한 단기 전력 수요 예측," 전기학회 논문지, Vol. 54A, No. 6, pp. 259-266, Jun., 2005.
- [23] 지평식, 남상천, 임재윤, 김정훈, "분류된 부하패턴을 근거로 한 단기 전력 수요 예측," 대한전기학회논문지, Vol. 47, No. 3, pp. 269-275, March, 1998.

저 자 소 개

**조 남 훈 (趙南熏)**



1992년 서울대 공대 전기공학과 졸업. 2000년 서울대 대학원 전기공학부 졸업(공박). 2000년~2001년 서울대 자동화시스템공동연구소 연구원. 2001년~2002년 삼성전자 DVS사업부 책임연구원. 2002년~현재 숭실대학교 전기공학부 조교수.  
Tel : 02-820-0643  
E-mail : nhjo@ssu.ac.kr  
Homepage : <http://ee.ssu.ac.kr/~nhjo>

**송 경 빈 (宋敬彬)**



1963년 9월 15일생. 1986년 연세대전기공학과 졸업. 1988년 동 대학원 전기공학과 졸업(석사). 1995년 텍사스 A&M전기공학과 졸업(공박). 1996년 한전전력연구원 선임연구원. 2002년 숭실대학교 전기제어시스템공학부 전임강사. 2004년~현재 숭실대학교 전기공학부 조교수. 2006년~현재 대한전기학회 전력기술부문회 편집위원장.  
Tel : 02-820-0648  
E-mail : kbsong@ssu.ac.kr

**노 영 수 (盧永秀)**



1961년 4월 1일생. 1984년 서울대 전기공학과 졸업. 1986년 서울대 대학원 전기공학과 졸업(석사). 1988년~1996년 한국전기연구원, 연구원. 2001년 미국 캘리포니아주립대(Davis) 응용과학과 졸업(박사)현재 숭실대학교 전기공학부 조교수  
Tel : 02-820-0663  
Fax : 02-817-7961  
E-mail : yroh@ssu.ac.kr

**강 대 승 (姜大昇)**



1966년 9월 16일생. 1988년 서울대 물리학과 졸업. 1990년 서울대 대학원 물리학과 졸업(석사). 1991년~1994년 금성사 안양연구소 연구원. 1999년 미국 Case Western Reserve 대학 물리학과 졸업(박사). 현재 숭실대학교 전기공학부 조교수  
Tel : 02-820-0642  
Fax : 02-817-7961  
E-mail : dkang@ssu.ac.kr